

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/009, 004 C  
Source: IFW16  
Date Processed by STIC: 06/20/2005

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 06/20/2005

PATENT APPLICATION: US/10/009,004C

TIME: 09:54:23

Input Set : A:\Revised Gregory Seq.txt

Output Set: N:\CRF4\06202005\J009004C.raw

```

3 <110> APPLICANT: Gregory, Richard L.
5 <120> TITLE OF INVENTION: Methods and Compositions for Controlling Dental Caries,
6   and Recombinant SmaA Polypeptides Useful for Same
8 <130> FILE REFERENCE: IU97
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/009,004C
C--> 11 <141> CURRENT FILING DATE: 2001-11-05
13 <150> PRIOR APPLICATION NUMBER: 60/132,312
14 <151> PRIOR FILING DATE: 1999-05-03
16 <160> NUMBER OF SEQ ID NOS: 6
18 <170> SOFTWARE: PatentIn Ver. 2.1
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 7
22 <212> TYPE: PRT
23 <213> ORGANISM: Streptococcus mutans
25 <400> SEQUENCE: 1
26 Glu Glu Gln Ser Gly Gly Thr
27   1               5
30 <210> SEQ ID NO: 2
31 <211> LENGTH: 7
32 <212> TYPE: PRT
33 <213> ORGANISM: Streptococcus mutans
35 <400> SEQUENCE: 2
36 Tyr Leu Met Lys Gly Gly Thr
37   1               5
40 <210> SEQ ID NO: 3
41 <211> LENGTH: 11
42 <212> TYPE: PRT
43 <213> ORGANISM: Streptococcus mutans
45 <400> SEQUENCE: 3
46 Met Ser Ser Gln Ala Lys Ala Asn Asn Ile Pro
47   1               5               10
50 <210> SEQ ID NO: 4
51 <211> LENGTH: 11
52 <212> TYPE: PRT
53 <213> ORGANISM: Streptococcus mutans
55 <220> FEATURE:
56 <221> NAME/KEY: MISC_FEATURE
57 <222> LOCATION:
58 <223> OTHER INFORMATION: Xaa is equivalent to any Amino Acid
60 <400> SEQUENCE: 4
W--> 61 Met Gln Arg Pro Thr Glu Phe Xaa Glu Asp Lys
62   1               5               10
65 <210> SEQ ID NO: 5

```

## RAW SEQUENCE LISTING

DATE: 06/20/2005

PATENT APPLICATION: US/10/009,004C

TIME: 09:54:23

Input Set : A:\Revised Gregory Seq.txt

Output Set: N:\CRF4\06202005\J009004C.raw

```

66 <211> LENGTH: 2985
67 <212> TYPE: DNA
68 <213> ORGANISM: Streptococcus mutans
70 <220> FEATURE:
71 <221> NAME/KEY: CDS
72 <222> LOCATION: (816)..(1820)
74 <400> SEQUENCE: 5
75 ggatcctcgc caccgagatc gtattgcctt tgtgctggatt gtctcaggtg aatttgaacg 60
77 cggcatgtct gtcaaccttg cccgcactgg taagagcggt aagctgtcaa acgtcactca 120
79 gtttatggca gaatctcgtg agaatgtaga aaatgctgtc gctggtgata ttatcggagt 180
81 ttacgataca ggaacttatc aggttggtga taccttaact gttggtaaaa ataaatttga 240
83 atttgagcca ctgccgacct ttacaccaga gctctttatg aaagtttctg ctaaaaatgt 300
85 tatgaagcag aagtcttttc ataaaggcat tgagcaattg gtgcaagaag gtgctataca 360
87 gctttatacc aactatcaaa ctggtgaata tatgcttgga gcagtcggtc agctccagtt 420
89 tgaagttttc aaacaccgca tggaaaatga atacaatgcg gaggttatca tgacacccat 480
91 gggtaagaaa acggtgcggt ggattaagga ggaagatctt gatgaacgta tgtcttccag 540
93 ccgcaatata ttggctaaag accgctttaa caagcccggt ttcctctttc aaaatgactt 600
95 ttctcttcat tggtttgcag ataaatatcc agatatagtt ttggaagaga agatgtaaca 660
97 atgatttggt taaatcttat aaataacgat caaccatttc ttggtttggt gtttttttat 720
99 atagtaaaaa tttttcaaaa aatatattac gtaagtattg ctaaatattt cttttgtggt 780
101 tcaatatagg tgaaaaaaga aaatgaagga agatt atg aat caa aaa ata gtc 833
102                                     Met Asn Gln Lys Ile Val
103                                     1           5
105 gtc att tcg tca ttt tac atg tta ggt gct cat tca ttt tca aag gca 881
106 Val Ile Ser Ser Phe Tyr Met Leu Gly Ala His Ser Phe Ser Lys Ala
107                                     10          15          20
109 gta tat cat aat gat agg agt gtg aaa ctt atg aaa aga att gat att 929
110 Val Tyr His Asn Asp Arg Ser Val Lys Leu Met Lys Arg Ile Asp Ile
111                                     25          30          35
113 aat cat caa gca caa cgt ttt tct att cgt aaa tat gca ttt gga gct 977
114 Asn His Gln Ala Gln Arg Phe Ser Ile Arg Lys Tyr Ala Phe Gly Ala
115                                     40          45          50
117 gca tct gtt tta att ggc tgt gtc ttt ttt cta ggt acc caa aat gtt 1025
118 Ala Ser Val Leu Ile Gly Cys Val Phe Phe Leu Gly Thr Gln Asn Val
119 55          60          65          70
121 tct gca caa gag cag gga act caa ttg cca gca agt gaa aac gca gtt 1073
122 Ser Ala Gln Glu Gln Gly Thr Gln Leu Pro Ala Ser Glu Asn Ala Val
123                                     75          80          85
125 gtg aac gtg gct gaa aat tca gtt gct atc agc caa gca gtt gca gat 1121
126 Val Asn Val Ala Glu Asn Ser Val Ala Ile Ser Gln Ala Val Ala Asp
127                                     90          95          100
129 aag gca gca act caa aca act cta aca gaa aca ccc caa gtt gaa gtt 1169
130 Lys Ala Ala Thr Gln Thr Thr Leu Thr Glu Thr Pro Gln Val Glu Val
131                                     105         110         115
133 gag gag aaa gaa agt aag gta aat gct cct gct tta aat gtc gat gac 1217
134 Glu Glu Lys Glu Ser Lys Val Asn Ala Pro Ala Leu Asn Val Asp Asp
135                                     120         125         130
137 aaa ggt gca aaa tcc aaa gaa gat gtg aac cct act att tca aag aca 1265
138 Lys Gly Ala Lys Ser Lys Glu Asp Val Asn Pro Thr Ile Ser Lys Thr

```

## RAW SEQUENCE LISTING

DATE: 06/20/2005

PATENT APPLICATION: US/10/009,004C

TIME: 09:54:23

Input Set : A:\Revised Gregory Seq.txt

Output Set: N:\CRF4\06202005\J009004C.raw

```

139 135          140          145          150
141 gca agt gaa gtg gaa gct tct gca gta act gct act gat act aaa aat 1313
142 Ala Ser Glu Val Glu Ala Ser Ala Val Thr Ala Thr Asp Thr Lys Asn
143          155          160          165
145 tca aat cca caa gtc aat gtt gaa act gac tca agt gaa aaa gac gaa 1361
146 Ser Asn Pro Gln Val Asn Val Glu Thr Asp Ser Ser Glu Lys Asp Glu
147          170          175          180
149 aat aaa atg gtc acc tcg gct cca gct aag gag act gag gca gaa caa 1409
150 Asn Lys Met Val Thr Ser Ala Pro Ala Lys Glu Thr Glu Ala Glu Gln
151          185          190          195
153 aat gag aaa gcg gtc aga gaa aat ctt atg caa aga caa gct aag gct 1457
154 Asn Glu Lys Ala Val Arg Glu Asn Leu Met Gln Arg Gln Ala Lys Ala
155          200          205          210
157 gtc tca att cca tcg caa ggc aat tat gtt ttc caa gaa aca act cct 1505
158 Val Ser Ile Pro Ser Gln Gly Asn Tyr Val Phe Gln Glu Thr Thr Pro
159 215          220          225          230
161 gta aaa aat gca gcc agt atg tcc agc cca acc caa ttt aac ttt gat 1553
162 Val Lys Asn Ala Ala Ser Met Ser Ser Pro Thr Gln Phe Asn Phe Asp
163          235          240          245
165 aaa gga gat aag gtt ttt tat gat aat gtt tta gaa gcg gat ggg cat 1601
166 Lys Gly Asp Lys Val Phe Tyr Asp Asn Val Leu Glu Ala Asp Gly His
167          250          255          260
169 caa tgg att agc tat gtg tct tac agt ggt att cgt cgc tat gct cct 1649
170 Gln Trp Ile Ser Tyr Val Ser Tyr Ser Gly Ile Arg Arg Tyr Ala Pro
171          265          270          275
173 att gct gtg aca att gaa gaa ttg aag caa aaa gaa att gtt cag caa 1697
174 Ile Ala Val Thr Ile Glu Glu Leu Lys Gln Lys Glu Ile Val Gln Gln
175          280          285          290
177 aat tta ccg gca caa gga acc tat cac ttt act aaa cag cag agc tta 1745
178 Asn Leu Pro Ala Gln Gly Thr Tyr His Phe Thr Lys Gln Gln Ser Leu
179 295          300          305          310
181 aaa atg aag cta aac tgt cta gtc cga ccc aat tct cgt ttt aca acg 1793
182 Lys Met Lys Leu Asn Cys Leu Val Arg Pro Asn Ser Arg Phe Thr Thr
183          315          320          325
185 gag atc acg ttt ttt atg ata agg ttt tagaagcgga tggacatcaa 1840
186 Glu Ile Thr Phe Phe Met Ile Arg Phe
187          330          335
189 tggattagct atgtgtccta cagtgggtatc cgtcgttatg ttgttatttg aaagcttacg 1900
191 acacaaccct ctccaattga aactaaagta tcaggtacta ttgtcatcca aaataaaacg 1960
193 gctcaacaat tcgatgttgt catttctaata gcttcaagca atcaaggcat aaaagaggta 2020
195 ttagtgccag tttggtcaga gcaaaacggg caggatgaca ttgtctggta tcaagcaact 2080
197 aaacaagggtg aaggcgttta taagggtgacc gttaagggtca gtgaccataa aaataatagc 2140
199 ggtaactatc atgtccatct ttattatctt ttggataatg gtgaacaaag aggagtcggg 2200
201 gcaacaatga ctgagggtgga agcaccagag cctgtagaaa caacaggtat cattagcatt 2260
203 gccataaaga gcagccaagg atttgatggt ttgattacta atgcttcag cactcaagac 2320
205 ataaaagagg ttttagtgcc ggtttggtca gaacaaaacg gacaggacga tattatttg 2380
207 tatcaagcaa ctaaacaaagg cgaaggcgtt tataagggtg ccgttaagggt cagtgaccat 2440
209 aaaaatgaca gtggtaacta taacattcac ctttattatc gccttgtaac tgggtgaatta 2500
211 aagggtgttg gaggaaaagac aacgacagta gaagccccta atagagtcaa tcttccagca 2560

```

## RAW SEQUENCE LISTING

DATE: 06/20/2005

PATENT APPLICATION: US/10/009,004C

TIME: 09:54:23

Input Set : A:\Revised Gregory Seq.txt

Output Set: N:\CRF4\06202005\J009004C.raw

```

213 caaggaactt atgttttcac taataaagtt gaggttaaaa atgaggccag aacatctagt 2620
215 ccaactcagt ttacctttaaa taaaggagaa agtatttact atgacagtat cttgaatgct 2680
217 gatggacatc aatggattag ctatcgttcc tacagtggta ttcgtcgtta tattatcatt 2740
219 ggttgaagta aaaaagggtta ggatgacaaa atcctgactt ttttgtgctt tagaattaat 2800
221 gttggataaa gtgtggagtt tgtgtctcgaa aaatagcagc gattgaatgt gtttataatt 2860
223 tgattcagac attagttttt atttcaagca aaaaatttga caaatcaaatt caattatatt 2920
225 acaatttttt aacgtatatt acaaaaatat atttgaaga tttattcaga tttggaggat 2980
227 ttatg 2985
230 <210> SEQ ID NO: 6
231 <211> LENGTH: 335
232 <212> TYPE: PRT
233 <213> ORGANISM: Streptococcus mutans
235 <400> SEQUENCE: 6
236 Met Asn Gln Lys Ile Val Val Ile Ser Ser Phe Tyr Met Leu Gly Ala
237 1 5 10 15
239 His Ser Phe Ser Lys Ala Val Tyr His Asn Asp Arg Ser Val Lys Leu
240 20 25 30
242 Met Lys Arg Ile Asp Ile Asn His Gln Ala Gln Arg Phe Ser Ile Arg
243 35 40 45
245 Lys Tyr Ala Phe Gly Ala Ala Ser Val Leu Ile Gly Cys Val Phe Phe
246 50 55 60
248 Leu Gly Thr Gln Asn Val Ser Ala Gln Glu Gln Gly Thr Gln Leu Pro
249 65 70 75 80
251 Ala Ser Glu Asn Ala Val Val Asn Val Ala Glu Asn Ser Val Ala Ile
252 85 90 95
254 Ser Gln Ala Val Ala Asp Lys Ala Ala Thr Gln Thr Thr Leu Thr Glu
255 100 105 110
257 Thr Pro Gln Val Glu Val Glu Glu Lys Glu Ser Lys Val Asn Ala Pro
258 115 120 125
260 Ala Leu Asn Val Asp Asp Lys Gly Ala Lys Ser Lys Glu Asp Val Asn
261 130 135 140
263 Pro Thr Ile Ser Lys Thr Ala Ser Glu Val Glu Ala Ser Ala Val Thr
264 145 150 155 160
266 Ala Thr Asp Thr Lys Asn Ser Asn Pro Gln Val Asn Val Glu Thr Asp
267 165 170 175
269 Ser Ser Glu Lys Asp Glu Asn Lys Met Val Thr Ser Ala Pro Ala Lys
270 180 185 190
272 Glu Thr Glu Ala Glu Gln Asn Glu Lys Ala Val Arg Glu Asn Leu Met
273 195 200 205
275 Gln Arg Gln Ala Lys Ala Val Ser Ile Pro Ser Gln Gly Asn Tyr Val
276 210 215 220
278 Phe Gln Glu Thr Thr Pro Val Lys Asn Ala Ala Ser Met Ser Ser Pro
279 225 230 235 240
281 Thr Gln Phe Asn Phe Asp Lys Gly Asp Lys Val Phe Tyr Asp Asn Val
282 245 250 255
284 Leu Glu Ala Asp Gly His Gln Trp Ile Ser Tyr Val Ser Tyr Ser Gly
285 260 265 270
287 Ile Arg Arg Tyr Ala Pro Ile Ala Val Thr Ile Glu Glu Leu Lys Gln
288 275 280 285

```

## RAW SEQUENCE LISTING

DATE: 06/20/2005

PATENT APPLICATION: US/10/009,004C

TIME: 09:54:23

Input Set : A:\Revised Gregory Seq.txt

Output Set: N:\CRF4\06202005\J009004C.raw

290	Lys	Glu	Ile	Val	Gln	Gln	Asn	Leu	Pro	Ala	Gln	Gly	Thr	Tyr	His	Phe
291		290					295					300				
293	Thr	Lys	Gln	Gln	Ser	Leu	Lys	Met	Lys	Leu	Asn	Cys	Leu	Val	Arg	Pro
294	305					310					315					320
296	Asn	Ser	Arg	Phe	Thr	Thr	Glu	Ile	Thr	Phe	Phe	Met	Ile	Arg	Phe	
297					325					330						335

RAW SEQUENCE LISTING ERROR SUMMARY      DATE: 06/20/2005  
PATENT APPLICATION: US/10/009,004C      TIME: 09:54:24

Input Set : A:\Revised Gregory Seq.txt  
Output Set: N:\CRF4\06202005\J009004C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; Xaa Pos. 8

**VERIFICATION SUMMARY**

DATE: 06/20/2005

PATENT APPLICATION: US/10/009,004C

TIME: 09:54:24

Input Set : A:\Revised Gregory Seq.txt

Output Set: N:\CRF4\06202005\J009004C.raw

L:10 M:270 C: Current Application Number differs, Replaced Application Number

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:61 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0